



**TURUN  
YLIOPISTO**  
Kauppakorkeakoulu

# **Ennakoivan analytiikan rooli finanssilaitosten riskienhallinnassa**

Tietojärjestelmätieteen  
Kandidaatintutkielma

Laatija:  
Olli Hukka

Ohjaaja:  
FT Kai kimppa

12.12.2024  
Turku

Turun yliopiston laatujärjestelmän mukaisesti tämän julkaisun alkuperäisyys on tarkastettu Turnitin OriginalityCheck -järjestelmällä.

Kandidaatintutkielma

**Oppiaine:** Tietojärjestelmätiede

**Tekijä:** Olli Hukka

**Otsikko:** Ennakoivan analytiikan rooli riskienhallinnassa finanssisektorilla

**Ohjaaja:** FT Kai Kimppa

**Sivumäärä:** 28 sivua

**Päivämäärä:** 12.12.2024

Tämä tutkielma tarkastelee ennakoivan analytiikan roolia finanssilaitosten riskienhallinnassa. Tutkielmassa vastataan kahteen keskeiseen kysymykseen: *Mitä on ennakoiva analytiikka?* ja *Miten sitä hyödynnetään finanssilaitosten riskienhallinnassa?* Tutkielma on kirjallisuuskatsaus, jossa analysoidaan ennakoivan analytiikan teoreettisia lähtökohtia, käytännön sovelluksia ja sen roolia finanssilaitosten riskienhallinnassa.

Ennakoiva analytiikka määritellään edistyneeksi analytiikan menetelmäksi, joka hyödyntää tilastollisia ja koneoppimistekniikoita ennustaakseen tulevia tapahtumia datan pohjalta. Tekniikoihin kuuluvat muun muassa regressioanalyysi, klusterointi ja päätöspuut, jotka mahdollistavat riskien ennakoivan tunnistamisen ja päätöksenteon optimoinnin. Tutkielma korostaa, että mallien tehokkuus riippuu merkittävästi datan laadusta ja valittujen analyysimenetelmien tarkoituksenmukaisuudesta.

Tutkielmassa perehdytään finanssilaitosten tyypillisimpiin riskeihin, jotka ovat: luottoriski, markkinariski, likviditeettiriski ja operatiivinen riski. Luottoriski, joka liittyy lainanottajien maksukyvyttömyyteen, tunnistetaan merkittävimäksi riskiksi. Markkinariski puolestaan koskee markkinoiden volatiliteetista aiheutuvia tappioita, kun taas likviditeettiriski liittyy lyhyen aikavälin maksuvelvoitteiden täyttämiseen. Operatiivinen riski kattaa sisäisten prosessien epäonnistumiset sekä ulkoiset häiriöt.

Tulokset osoittavat, että ennakoiva analytiikka tehostaa riskienhallintaa tarjoamalla tarkkoja näkemyksiä esimerkiksi luottokelpoisuudesta, markkinatrendeistä, kassavirran vaihteluista ja operatiivisista haavoittuvuuksista. Lisäksi tutkielma suosittelee kehittyneiden teknologioiden, kuten selitettävän tekoälyn, integrointia mallien läpinäkyvyyden ja sääntelyn noudattamisen varmistamiseksi.

Tutkielman johtopäätöksenä on, että ennakoivasta analytiikasta on tullut olennainen osa finanssialan riskienhallintastrategioita. Ennakoivan analytiikan avulla instituutiot voivat vähentää riskejä, tehdä tietoon perustuvia päätöksiä ja lisätä vakautta sekä sidosryhmien luottamusta.

**Avainsanat:** ennakoiva analytiikka, riskienhallinta, finanssilaitos

# SISÄLLYS

<b>1</b>	<b>Johdanto</b>	<b>6</b>
1.1	Johdatus aihepiiriin	6
1.2	Tutkielman tavoite, rakenne, rajaukset ja menetelmät	6
<b>2</b>	<b>Ennakoiva analytiikka</b>	<b>8</b>
2.1	Määritelmä ennakoivalle analytiikalle ja käytännön esimerkkejä	8
2.2	Ennakoivan analytiikan tekniikat	9
2.3	Datan merkitys ennakoivan analytiikan järjestelmissä	12
2.4	Esimerkki ennustavasta mallista	13
<b>3</b>	<b>Ennakoivan analytiikan käyttö finanssilaitosten riskienhallinnassa</b>	<b>16</b>
3.1	Tyypillisiä finanssilaitosten riskejä	16
3.2	Ennakoiva analytiikka finanssilaitosten riskienhallinnan apuvälineenä	19
<b>4</b>	<b>Yhteenveto ja johtopäätökset</b>	<b>24</b>
	<b>Lähteet</b>	<b>26</b>

## **KUVIOT**

Kuva 1 Yksinkertainen luokittelupuumalli (mukaillen Fawcett & Provost 2013)	11
Kuva 2 Ennustavan mallin rakentamisen prosessi (mukaillen Nyce, 2007; Fawcett & Provost, 2013)	15
Kuva 3 Finanssilaitoksen tyypillisiä riskejä (mukaillen Leo ym., 2019)	16

## **TAULUKOT**

Taulukko 1 Tieteellinen viitekehys ennakoivan analytiikan käytöstä finanssilaitosten riskienhallinnassa	24
---	----

# 1 Johdanto

## 1.1 Johdatus aihepiiriin

Yksi Yhdysvaltojen suurimmista finanssilaitoksista, Lehman Brothers, ajautui konkurssiin syyskuussa 2008. Konkurssia pidetään merkittävänä tapahtumana, sillä sen katsotaan kiihdyttäneen Yhdysvaltojen asuntomarkkinoiden subprime-lainakriisin laajentumista globaaliksi finanssikriisiksi. (STT, 2018.) Kriisin juurisyinä pidetään holtitonta lainanantoa ja välinpitämätöntä riskienhallintaa (Forbes, 2023).

Globaalin finanssikriisin jälkeen finanssilaitosten riskienhallinnasta on tullut entistä keskeisempi osa toimialaa (Leo ym., 2019). Perinteisiin tilastollisiin malleihin perustuvia riskienhallintajärjestelmiä on korvattu uusilla, muun muassa koneoppimista hyödyntävillä, malleilla, jotka kykenevät käsittelemään suurempia määriä dataa. Suuret datamäärät yhdessä kehittyneiden koneoppimisalgoritmien kanssa antavat yrityksille mahdollisuuden tarkempaan riskikartoitukseen reaaliajassa. (Olaiya ym., 2024.) Tämä niin sanottu ennakoiva analytiikka pitää sisällään joukon edistyneitä analytiikan tekniikoita kuten regressio, klusterointi, luokittelupuut (engl. classification tree) ja koneoppiminen (engl. machine learning, ML) (Tummino, 2018).

Finanssilaitokset koetaan yleisesti konservatiivisiksi ja kömpelöiksi, kun puhutaan uusien teknologioiden implementoinnin kontekstissa. Ajatus juontaa juurensa todennäköisesti finanssilaitosten laajasta regulaatiosta, mikä voi estää niitä implementoimasta uusimpia järjestelmiä. Regulaatiosta huolimatta pirstaloituneet ja yhä dynaamisemmat markkinat pakottavat instituutiot löytämään tehokkaampia keinoja päätöksenteon tueksi (Olaiya ym., 2024). Ensi sijassa päätöksenteon tueksi halutaan uusinta mahdollista dataa, jota voidaan jalostaa ajankohtaiseksi informaatioksi kehittyneissä analytiikan järjestelmissä (Javaid, 2024).

## 1.2 Tutkielman tavoite, rakenne, rajaukset ja menetelmät

Tutkielmassa selvitetään, miten finanssilaitokset hyödyntävät ennakoivaa analytiikkaa riskienhallinnassa.

Tutkimuskysymykset ovat seuraavat:

1. Mitä on ennakoivan analytiikka?
2. Millainen rooli ennakoivalla analytiikalla on finanssilaitosten riskienhallinnassa?

Tutkielman toisessa luvussa vastataan ensimmäiseen tutkimuskysymykseen. Luvussa määritetään ennakoivan analytiikan käsite ja perehdytään muutamiin oleellisiin ennakoivan analytiikan tekniikoihin. Lisäksi siinä käsitellään datan merkitystä ennakoivan analytiikan järjestelmissä. Lopuksi luvussa käsitellyt teemat kootaan yhteen ja lukijalle havainnollistetaan ennustavan mallin rakentamisen prosessi. Toisen luvun jälkeen lukija ymmärtää mitä ennakoivalla analytiikalla tarkoitetaan ja miten ennustava malli rakennetaan.

Tutkielman kolmas luku vastaa toiseen tutkimuskysymykseen ja se aloitetaan perehtymällä tyypillisiin finanssilaitosten riskeihin. Tämän jälkeen keskitytään ennakoivan analytiikan rooliin niiden hallinnassa. Tutkielmassa ei käydä läpi kaikkia finanssilaitosten riskejä, sillä se kokonaisuus voisi itsessään toimia kandidaatintutkielman aiheena. Tarkastelu rajautuu yleisimpiin ja merkittävimpiin riskeihin. Neljäs luku sisältää tutkielman yhteenvedon ja johtopäätökset.

Tutkielmassa keskitytään vain finanssilaitoksiin ja niiden riskienhallintaan. Finanssilaitoksella tässä tutkielmassa tarkoitetaan pankkeja, vakuutus- ja sijoituspalveluyhtiöitä sekä muita liiketoimintaa harjoittavia finanssialan toimijoita. Tutkimuskysymysten kannalta ei ole tarpeellista erotella näitä toisistaan. Lisäksi ennakoivan analytiikan teknisiin perusteisiin ei ole tarvetta pureutua erityisen syväälle, koska työ on tietojärjestelmätieteen alalta. Oleellista on hahmottaa mitä sillä tarkoitetaan ja kuinka sitä hyödynnetään erilaisten riskityyppien hallinnassa.

Tutkielma toteutettiin kirjallisuuskatsauksena. Aihepiiriin sopivaa kirjallisuutta haettiin Scopus-, Volter- ja Google Scholar-tietokannoista. Lähdemateriaalin hakuun käytettiin sopivia avainsanoja ja niiden yhdistelmiä. Oleellisimmat hakutermit olivat ”riskienhallinta”, ”ennakoiva analytiikka” ja ”rahoitus”. Ennakoivan analytiikan nopean kehityksen takia lähdemateriaalina pyrittiin käyttämään mahdollisimman tuoreita tutkimuksia.

## 2 Ennakoiva analytiikka

Tutkielman ensimmäisessä sisältöluvussa kartoitetaan ennakoivan analytiikan perusluonne. Kartoitus on luonnollista aloittaa käsitteen määritelmällä, jonka jälkeen perehdytään ennakoivan analytiikan sovellusalueisiin. Tämän jälkeen keskitytään ennakoivan analytiikan tekniikoihin ja kehitykseen. Tekniikoita käsitellään ylempällä käytännön tasolla syventymättä liikaa niiden teknisiin perusteisiin. Koska ennakoiva analytiikka perustuu dataan, luvun viimeisessä alaluvussa käsitellään laajasti datan merkitystä ja mahdollisia haasteita, mitä se voi aiheuttaa, kun instituutio päättää ottaa käyttöön ennakoivan analytiikan järjestelmiä.

### 2.1 Määritelmä ennakoivalle analytiikalle ja käytännön esimerkkejä

Ennakoiva analytiikka pyrkii vastaamaan kysymykseen, “Mitä tulee todennäköisesti tapahtumaan?”. Se kattaa laajan kirjon erilaisia tilastollisia tekniikoita perinteisistä ennustavista malleista koneoppimiseen ja tiedonlouhintaan. Perustana on instituution keräämä, tai kolmansilta osapuolilta hankittu, data, jota analysoimalla voidaan tehdä ennustuksia tulevista tapahtumista. (Ravi ym., 2018.) Ennakoivan analytiikan määrittelystä tutkijat ovat hyvin yksimielisiä. Lähes poikkeuksetta määritelmässä toistuu historiadata ja sen muokkaaminen ennustavien mallien avulla informaatioksi, joka auttaa päätöksentekijää hahmottamaan tulevaisuuden tapahtumia. Ennakoivaa analytiikkaa pidetään yhtenä tämän sukupolven edistyneimpänä teknisenä innovaationa yhdessä oppimisalgoritmien ja koneoppimisen kanssa (Močarníková & Greguš, 2020).

Erilaiset instituutiot hyödyntävät ennakoivaa analytiikkaa useista syistä. Sitä hyödynnetään esimerkiksi päätöksenteon tukena, riskienhallinnassa ja resurssien optimoinnissa. Sen avulla organisaatiot voivat tehdä päätöksiä, jotka perustuvat tietoon intuition tai aiempien kokemusten sijaan. (Fawcett & Provost, 2013, s. 5.) Analysoimalla historiallisia tietoja ennakoivan analytiikan avulla voidaan ennustaa tulevia tuloksia ja trendejä. Tämä auttaa yrityksiä ennakoimaan haasteita ja hyödyntämään mahdollisuuksia. (Javaid 2024.) Fawcettin ja Provostin (2013, s. 6) mukaan dataan perustuva päätöksenteko tarjoaa merkittävän edun instituutioille, kun niiden tulee ennustaa ja suunnitella tulevia tapahtumia.

Ennakoivaa analytiikkaa hyödynnetään laajalti riskienhallinnassa. Ennustavien mallien avulla voidaan arvioida erilaisten riskien, kuten lainojen laiminlyöntien, todennäköisyyttä. Niiden avulla instituutioilla on mahdollisuus reagoida ja varautua riskeihin ennakoivasti. (Nyce, 2007.)

Finanssisektorilla ennakoivaa analytiikkaa käytetään laajalti luottoriskien arviointiin, petosten havaitsemiseen ja osakekurssien muutosten ennustamiseen (Fawcett & Provost, 2013, s. 7).

Liiketoimintaa harjoittavat instituutiot voivat hyödyntää ennakoivaa analytiikkaa resurssien ja prosessien optimoinnissa. Esimerkiksi toimitusketjun hallinnassa ennustavien mallien avulla voidaan ennustaa kysyntää ja optimoida varastonhallintaa (Chopra ym., 2019). Kyky optimoida resursseja on kriittinen toimialoilla, joilla tehokkuus vaikuttaa suoraan kannattavuuteen.

## 2.2 Ennakoivan analytiikan tekniikat

Ennakoivan analytiikan tekniikat vaihtelevat ongelman luonteesta ja käytettävissä olevan datan tyypistä riippuen aina perinteisistä tilastollisista malleista kehittyneempiin koneoppimisalgoritmeihin. Tämä alaluku käsittelee keskeisempiä ennakoivan analytiikan tekniikoita: regressioanalyysiä, klusterointia, luokittelupuita ja koneoppimista, sekä näiden kehitystä ja kehityksen mukana tulleita haasteita.

Ennakoivan analytiikan juuret ylettyvät 1900-luvulle, mutta suurimmat ja merkittävimmät läpimurrot tehtiin 2000-luvulla (Močarníková & Greguš, 2020). Uusien tekniikoiden kehitykseen vaikutti instituutioiden halu hyödyntää massadataa päätöksenteossa (Kalaian ym., 2017). Nykyään dataan perustuva päätöksenteko on normaali osa instituutioiden prosesseja. 2010-luvulla merkittävimpiä kehitysaskelaita olivat kone- ja syväoppimisalgoritmien laajempi hyödyntäminen (Farayola ym., 2024). Kyseiset algoritmit pystyvät käsittelemään suurempia määriä monimutkaisempaa dataa. Adebukolan ym. (2022) mukaan syväoppimisalgoritmipohjaiset ennakoivan analytiikan järjestelmät voivat antaa tarkempia tuloksia tehokkaammin.

Regressioanalyysi on yksi ennakoivan analytiikan perustavista menetelmistä. Lineaarinen regressio auttaa ymmärtämään riippuvien ja riippumattomien muuttujien välisiä suhteita. Se sopii erityisesti trendien ennustamiseen. (Javaid, 2024.) Regressiotekniikat ovat jo pitkään olleet keskeisessä roolissa liiketoimintaa koskevassa päätöksenteossa, sillä ne tarjoavat helposti tulkittavia malleja, joiden käyttöönotto on vaivatonta (Shmueli & Koppius, 2011). Tätä näkemystä tukevat myös Bücker ym. (2022), jotka korostavat, että tekoälyn kehityksestä huolimatta regressio on edelleen varteenotettava ennakoivan analytiikan tekniikka sen yksinkertaisuuden ja läpinäkyvyyden vuoksi. Esimerkiksi finanssisektorin laaja regulaatio asettaa haasteita edistyneempien mallien käyttöönottoon niiden musta laatikko luonteensa takia. Mustalla laatikolla viitataan malliin, jonka toimintaperiaatetta on vaikea ymmärtää. Siksi esimerkiksi logistista regressiota käytetään edelleen laajalti luottojen pisteytyksessä. (Bücker ym., 2022.)

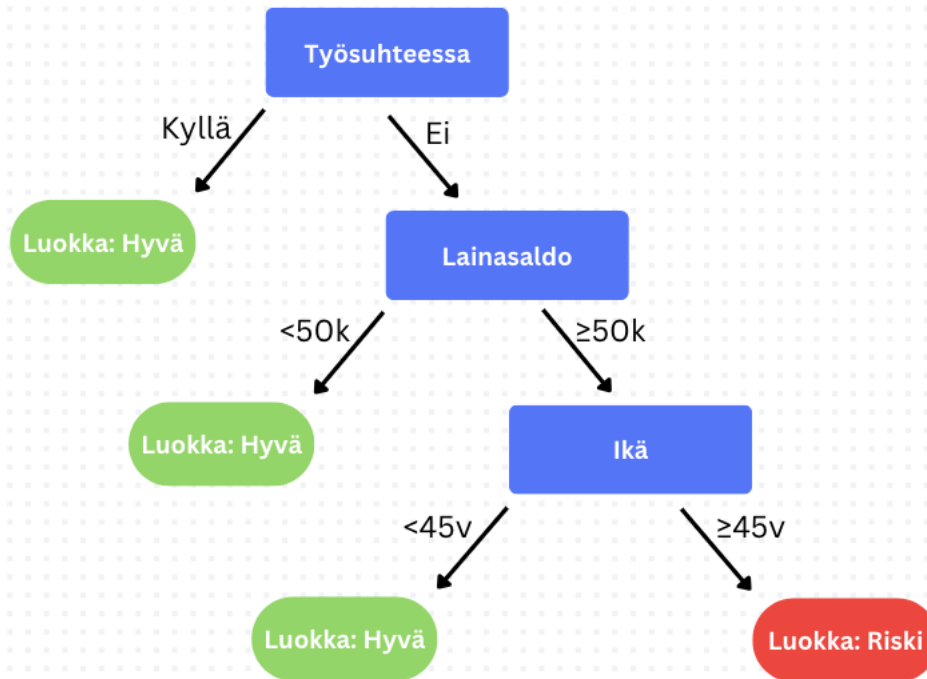
Regressio sopii tehtäviin, joissa pitää selvittää kuinka paljon tai kuinka todennäköisesti jokin asia tapahtuu. Tästä syystä sen tuottamat tulokset ovat numeerisia. Regression lisäksi ennakoiva

analytiikka pitää sisällään datan luokitteluun soveltuvia tekniikoita, kuten klusterointi ja luokittelupuut, jotka tuottavat kategorisia, yleensä binäärisiä, tuloksia. Klusteroinnin tavoitteena on segmentoida muuttujia ryhmiin tietyn tuloksen ennustamisen sijaan. (Fawcett & Provost, 2013, s. 21-26.) Se on oleellinen osa ennakoivaa analytiikkaa, vaikkei itsessään tarjoa vastauksia tulevaisuuden kysymyksiin. Klusteroinnilla voidaan esimerkiksi ryhmitellä samankaltaisia asiakkaita ryhmiin ja kohdentaa mainontaa (Tummino, 2018). Luokittelutekniikoita käytetään yleensä yhdessä muiden tekniikoiden kanssa, jolloin analyysi voidaan kohdentaa tietylle kohderyhmälle (Shmueli & Koppius, 2011).

Toinen luokitteluun sopiva tekniikka on luokittelupuun. Se tarjoaa intuitiivisen tavan tehdä dataan perustuvia päätöksiä ja sen visuaalinen luonne tekee siitä erityisen houkuttelevan ei-teknisille sidosryhmille (Salzberg, 1994). Kuva 1 havainnollistaa luokittelupuun perusidean. Luokittelupuun jokainen haara sisältää säännön, esimerkiksi henkilö on työsuhteessa, ja jokainen oksa edustaa tiettyä tulosta, esimerkiksi kyllä tai ei. Muuttuja etenee ominaisuuksiensa mukaan haarasta toiseen ja lopputuloksena muuttuja voidaan kategorisoida ominaisuuksiensa perusteella. Puumallista johtuen jokainen datapiste voi päätyä vain yhteen lehteen, joka edustaa muuttujan luokkaa. Luokittelupuumalleja käytetään luokittelun lisäksi myös regressiotöhtävissä. Tehokkuus perustuu kykyyn analysoida useita muuttujan ominaisuuksia kerralla. Luokittelupuumallit ovat suosittuja, koska ne ovat helposti ymmärrettäviä, selitettäviä ja toteutettavia ratkaisuja. (Fawcett & Provost, 2013, s. 63-64.)

Kuvan 1 yksinkertainen luokittelupuumalli havainnollistaa, kuinka finanssilaitokset luokittelevat asiakkaitaan lainanannon yhteydessä. Se on mukaelma Fawcettin ja Provostin (2013) esittämästä mallista. Esimerkki on yksinkertaistettu, jotta lukijan helppo ymmärtää mallin idea ja tarkoitus. Lisäksi voidaan olettaa, että finanssilaitoksella on tiedossa tarvittavat tiedot asiakkaan riskiluokan arvioimiseksi. Ensimmäinen lainan takaisinmaksun kannalta merkittävä tekijä on asiakkaan työllisyystilanne. Jos hän on työsuhteessa, hänen riskiluokkansa on hyvä ja lainan laiminlyönnin todennäköisyys on pieni. Jos hän ei ole työsuhteessa, finanssilaitoksen tulee tarkastella muita muuttujia. Seuraava luokittelusääntö liittyy lainasaldoon. Jos lainasaldo on tarpeeksi pieni niin laiminlyönnin todennäköisyys pienenee. Kuten tutkielmassa aiemmin todettiin, luokittelupuut tarjoavat intuitiivisen tavan tehdä dataan perustuvia päätöksiä. Jotta mallit toimisivat mahdollisimman tehokkaasti, finanssilaitosten tulisi tunnistaa merkittävimmät asiakkaan ominaisuudet, jotka vaikuttavat mahdolliseen lainan laiminlyöntiin.

**Yksinkertainen luokittelupuumalli, jolla voidaan arvioida lainan laiminlyönnin todennäköisyyttä**



Kuva 1 Yksinkertainen luokittelupuumalli (mukaillen Fawcett & Provost, 2013)

Koneoppiminen on viime vuosina laajentanut merkittävästi ennakoivan analytiikan alaa. Tekniikat, kuten neuroverkot ja tukivektorikoneet, mahdollistavat monimutkaisempien epälineaaristen relaatioiden analysoinnin, joihin perinteiset tilastolliset mallit eivät pysty (Goodfellow ym., 2016). Lisäksi koneoppimisalgoritmit pystyvät käsittelemään valtavia määriä dataa perinteisiin tilastollisiin malleihin verrattuna, mikä johtaa tarkempiin ennustuksiin (Adebukola ym., 2022). Nämä mallit oppivat tarkemmiksi ilman manuaalista ohjelmointia ja paranevat ajan myötä (Javaid, 2024).

Koneoppimis pohjaiset ennustavat mallit asettavat haasteita ennustavan voiman ja mallin selitettävyyden välille (Bücker ym., 2022). Kehittyneiden algoritmien toimintaperiaatteita on vaikea ymmärtää, mikä voi olla haaste laajan regulaation aloilla, kuten finanssisektorilla, missä myös ennustuksien perusteita tulee ymmärtää. Toisaalta Javaid (2024) kertoo koneoppimisen mullistaneen juuri finanssisektorin. Hän korostaa koneoppimis pohjaisten ennustavien mallien hyviä puolia, kuten kykyä hyödyntää erilaisia datatyyppejä sekä kykyä hyödyntää ajallisesti uutta dataa.

Ennustavien mallien selitettävyyttä voidaan parantaa selitettävän tekoälyn (engl. explainable artificial intelligence, XAI) avulla. Selitettävä tekoäly pyrkii avaamaan mallin käyttäjälle, miten tietty tulos on saatu ja mitkä muuttujat ovat vaikuttaneet tulokseen. (Bücker ym., 2022.)

Selitettävien mallien kehittäminen auttaa instituutioita rakentamaan asiakkaiden luottamusta tekoälyjärjestelmiin. Kehityksessä tulee huomioida mallien oikeudenmukaisuus ja puolueettomuus, jotta voidaan varmistaa tasapuoliset tulokset eri väestöryhmien välillä. (Farayola ym., 2024.)

Ennakoivan analytiikan menetelmät kehittyvät jatkuvasti teknologisten innovaatioiden ja datan saatavuuden kasvun seurauksena. Tasapaino mallien läpinäkyvyyden ja ennustavan voiman välillä on edelleen haaste, varsinkin kun organisaatiot pyrkivät implementoimaan entistä enemmän tekoälypohjaisia malleja (Farayola ym., 2024). Mallien selitettävyyden lisäksi datan hallinta ja prosessointi aiheuttaa haasteita (Nyce, 2007). Fawcettin ja Provostin (2013, s. 16) mielestä instituutioiden tulisi nähdä data omaisuutena, jonka jälkeen voidaan pohtia, kuinka paljon siihen halutaan investoida. Lopuksi on tärkeää huomauttaa, että ennakoivan analytiikan järjestelmien implementoiminen vaatii kehittyneen teknologian lisäksi myös organisaatiotason linjauksia, tietojen lukutaitoa, sekä selkeää viestintää teknisten ja ei-teknisten sidosryhmien välillä, jotta investoinnista saadaan maksimaalinen hyöty (Chopra, 2019).

### **2.3 Datan merkitys ennakoivan analytiikan järjestelmissä**

Ennakoivan analytiikan järjestelmien toimivuus perustuu valtaviin määriin dataa (Javaid, 2024). Suuremmat määrät dataa mahdollistaa heikompien suhteiden ja trendien havaitsemisen. Fawcett ja Provost (2013, s. 8) väittävät, että moderni ennakoiva analytiikka on riippuvaista massadatan räjähdysmäisestä kasvusta. Tummino (2022) korostaa, että erityisesti finanssisektori on hyötynyt datan saatavuuden kasvusta. Tämä alaluku käsittelee datan merkitystä ennakoivan analytiikan järjestelmissä. Lisäksi siinä nostetaan esille merkittäviä ennakoivaan analytiikkaan liittyviä haasteita ja tulevaisuuden mahdollisuuksia, jotka tulisi tunnistaa ennen käyttöönottoa.

Datan määrän ja etenkin saatavuuden kasvu tulee vahvistamaan ennakoivan analytiikan järjestelmiä. Määrän ja saatavuuden sijaan instituutiot voisivat keskittyä enemmän käytettävän datan laatuun, joka voi vaihdella paljon. Tietoja voi puuttua, data voi olla puolueellista tai siihen on voinut muodostua vinoumia. Kaikki edellä mainitut ongelmat voivat heikentää ennustavan mallin suorituskykyä. (Nyce, 2007.) Ongelmat voidaan kuitenkin välttää esikäsittelemällä käytettävä data. Esikäsitteily voi pitää sisällään puuttuvien tietojen manuaalista lisäystä, tietokantojen yhtenäistämistä ja ääripäiden tasoittelua. (Javaid, 2024.) Puhtaan ja valmistellun datan merkitys on suuri, sillä datassa olevat virheet voivat johtaa virheellisiin riskiarvioihin ja huonoihin taloudellisiin

päätöksiin (Tummino, 2018). Datan esikäsittely on kuitenkin aikaa vievä ja kallis prosessi varsinkin, jos instituutio ei ole luontaisesti dataorientoitunut. Lisäksi instituutioilla voi olla käytössä datan varastointiin vanhoja perinnejärjestelmiä, jotka eivät ole integroitavissa moderneihin ennakoivan analytiikan järjestelmiin. (Nyce, 2007.) Ennakoivan analytiikan järjestelmän käyttöönotto voi vaatia instituutiolta laajoja muutoksia ja prosessien uudelleen järjestelyä.

Koneoppimisen kehityksen myötä instituutiot pyrkivät jatkuvasti hyödyntämään ajallisesti uutta dataa parantaakseen ennustavia mallejaan (Javaid, 2024). Kyky analysoida mahdollisimman reaaliaikaista dataa antaa instituutioille mahdollisuuden tehdä tietoon perustuvia päätöksiä nopeammin (Nyce, 2007). Reaaliaikainen tieto on välttämätöntä nopeasti muuttuvilla markkinoilla, kun halutaan ymmärtää äkillisiä muutoksia (Javaid, 2024).

Data voidaan jakaa strukturoituun ja strukturoimattomaan dataan. Strukturoitu data sisältää siististi järjesteltyjä tietokantoja, kuten tapahtumahistorioita ja taseita, kun taas strukturoimaton data koostuu vähemmän organisoiduista formaateista, kuten sähköposteista, sosiaalisen median julkaisuista tai äänitallenteista. (Zhang ym., 2020.) Koneoppiminen on mahdollistanut strukturoimattoman datan hyödyntämisen ennakoivan analytiikan järjestelmissä luonnollisen kielen käsittelyn avulla (engl. natural language processing, NLP). Kehittyneiden NLP-tekniikoiden avulla instituutiot voivat hyödyntää esimerkiksi asiakasarvosteluja ja sosiaalisen median julkaisuja ennakoivan analytiikan järjestelmissä. (Javaid, 2024.) Kun instituutiot oppivat keräämään ja analysoimaan erilaisia datatyyppejä, yksityisyyden suojaa koskevat kysymykset nousevat esille. Esimerkiksi terveydenhuollossa käsitellään jatkuvasti arkaluontoisia potilastietoja, joiden hallinnassa tulee noudattaa erityistä varovaisuutta. (Adebukola ym., 2022.)

Johtopäätöksenä voidaan todeta, että data on ennakoivan analytiikan elinehto, koska ilman dataa ennustavia malleja ei voi rakentaa. Laadukkaan, monipuolisen ja reaaliaikaisen datan avulla organisaatiot voivat hyödyntää ennakoivaa analytiikkaa hallitakseen riskejä ja parantaakseen päätöksentekoa (Javaid, 2024). Farayola ym. (2024) huomauttavat että instituutioiden on mallien tarkkuuden lisäksi varmistettava niiden läpinäkyvyys ja oikeudenmukaisuus. He korostavat, että datan käsittelyyn liittyvien prosessien tulee olla linjassa eettisten standardien kanssa.

## **2.4 Esimerkki ennustavasta mallista**

Tutkielmassa on tähän mennessä käsitelty ennakoivan analytiikan tekniikoita ja perehdytty datan merkitykseen ennakoivan analytiikan järjestelmissä. Tässä alaluvussa sovelletaan aiemmin käsiteltyjä teemoja ja esitetään havainnollistava esimerkki ennustavan mallin rakentamisen

prosessista. Kuvio 2 havainnollistaa prosessin. Esimerkissä käytettävä malli rakennetaan hyödyntäen valvottua oppimista (engl. supervised learning), joka on yksi koneoppimisen muoto.

Valvottu oppiminen on koneoppimisen muoto, jossa algoritmi muodostaa mallin opetusdatan avulla. Opetusdatassa kaikkien muuttujien ja tulosten tulee olla ennalta määrättyjä, jotta algoritmi oppii syy-seuraussuhteet selittävien muuttujien ja tuloksen välillä. (Farayola ym., 2024.) Esimerkki on mukaelma Shmuelin ja Koppiuksen (2011) sekä Fawcettin ja Provostin (2013) esittämistä prosesseista.

Ennustavat mallit perustuvat tietyn yksittäisen kohdemuuttujan (engl. target variable) arvon selvittämiseen. Muuttujan arvo voi olla numeerinen, binäärinen, todennäköisyys tai yhdistelmä näistä. Prosessi aloitetaan määrittelemällä tavoite, joka määrää millainen kohdemuuttuja halutaan selvittää ja minkä tyyppisen arvon sen tulisi saada. (Fawcett & Provost, 2013, s. 44–50; Shmueli & Koppius, 2011.) Esimerkiksi finanssilaitos voisi haluta selvittää, jättääkö asiakas lainan maksamatta. Tällöin kohdemuuttuja olisi binäärinen, joka voi saada arvokseen kyllä tai ei. Lisäksi tulee selvittää mitkä tiedossa olevat muuttujat (selittävät muuttujat) vaikuttavat kohdemuuttujan arvoon (Fawcett & Provost, 2013, s. 48-50). Lainan takaisin maksuun voisi vaikuttaa muun muassa velallisen ikä, ammatti ja omaisuus. Selkeän ja laskettavissa olevan tavoitteen määrittäminen on erittäin oleellinen osa ennustavan mallin rakentamisen prosessia, sillä se määrittelee prosessin seuraavat vaiheet (Shmueli & Koppius, 2011).

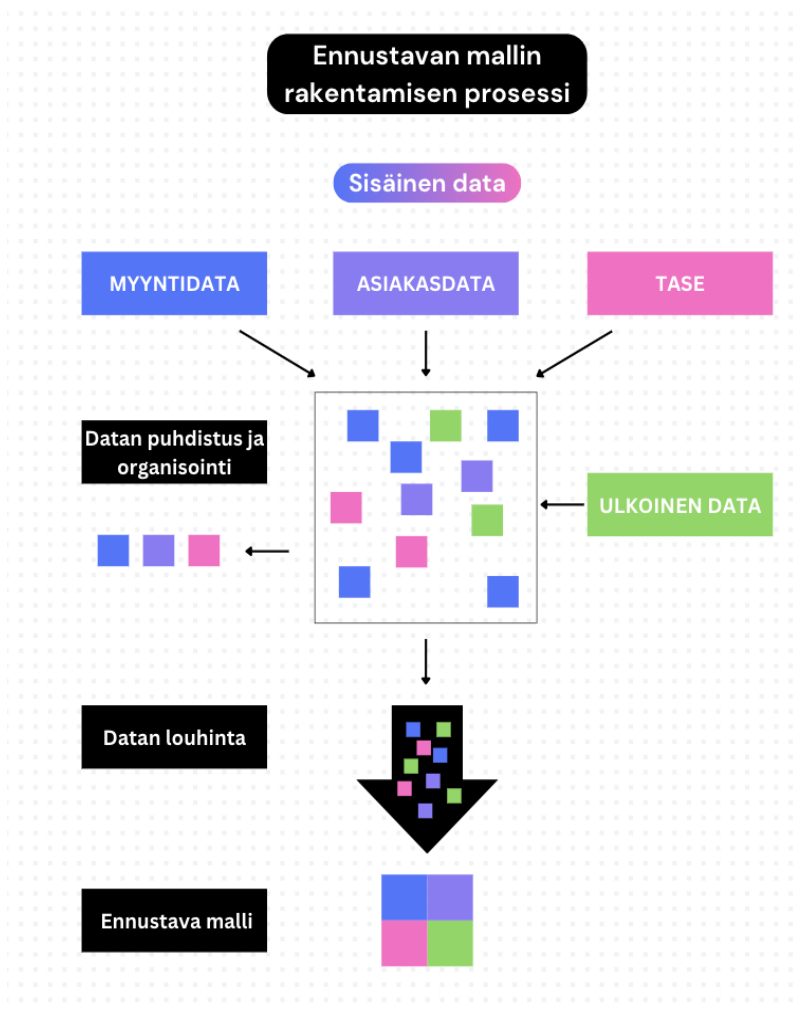
Kun mallin rakentajalla on tiedossa kohdemuuttuja, seuraavaksi kerätään kaikki mahdolliset tiedot, joka voivat vaikuttaa sen arvoon. Dataa kerätään sekä sisäisistä että ulkoisista lähteistä. Tavoitteena on löytää kaikki mahdolliset ja relevantit selittävät muuttujat. Keräyksen jälkeen data puhdistetaan ja valmistellaan prosessin seuraavaa vaihetta varten. Puhdistus voi sisältää ääripäiden tasoittelua, puuttuvien tietojen syöttämistä ja tarpeettomien tietojen poistamista. (Fawcett & Provost, 2013, s. 28–39; Shmueli & Koppius, 2011.)

Data täytyy muokata sopivaan muotoon, jotta sitä voidaan käyttää mallin opetukseen. Datalähtöiset algoritmit perustuvat tiedonlouhintatekniikoihin, kuten neuroverkkoihin, jotka löytävät opetusdatasta trendejä, segmenttejä ja syy-seuraussuhteita ilman manuaalista ohjelmointia. (Javaid, 2024; Shmueli & Koppius, 2011.)

Tiedonlouhinta pitää sisällään koneoppimisen ja neuroverkkojen lisäksi useita erilaisia tekniikoita, joihin tässä tutkielmassa ei perehdytä. Tärkeää on ymmärtää louhinnan perusidea, joka on löytää valtavista datamassoista selittäviä tekijöitä kohdemuuttujaan arvoon (Fawcett & Provost, 2013,

s. 40). Fawcettin ja Provostin (2013, s. 25-26) mukaan on kuitenkin oleellista erottaa milloin kyseessä on tiedonlouhinta ja milloin sen tulosten hyödyntäminen. Tiedonlouhinnalla pyritään löytämään relevanttia informaatiota datasta ja rakentamaan malli ennakoivan analytiikan tarkoituksiin. Sen tulosten hyödyntämisessä tiedonlouhinnalla opetettua mallia käytetään uuden datan analysointiin, jossa kohdemuuttujat ovat tuntemattomia. (Fawcett & Provostin, 2013, s. 25-26.) Kun algoritmi on opetettu ennustava malli tulee testata ja hienosäätää parhaiden tulosten saavuttamiseksi (Shmueli & Koppius, 2011).

Lopuksi ennustavaa mallia testataan testidatalla, jonka tulee poiketa opetuksessa käytetystä datasta. Tällöin voidaan varmistaa, että malli on oikeasti geneerinen eikä ylisovittunut opetusdataan. Kun mallin suorituskykyyn ollaan tyytyväisiä, se voidaan implementoida instituution tietojärjestelmiin. Lopuksi on hyvä huomauttaa, että ennustavien mallien suorituskykyä tulee arvioida tasaisin väliajoin ja varmistaa mallien käyttämän datan ajantasaisuus. (Shmueli & Koppius, 2011.)

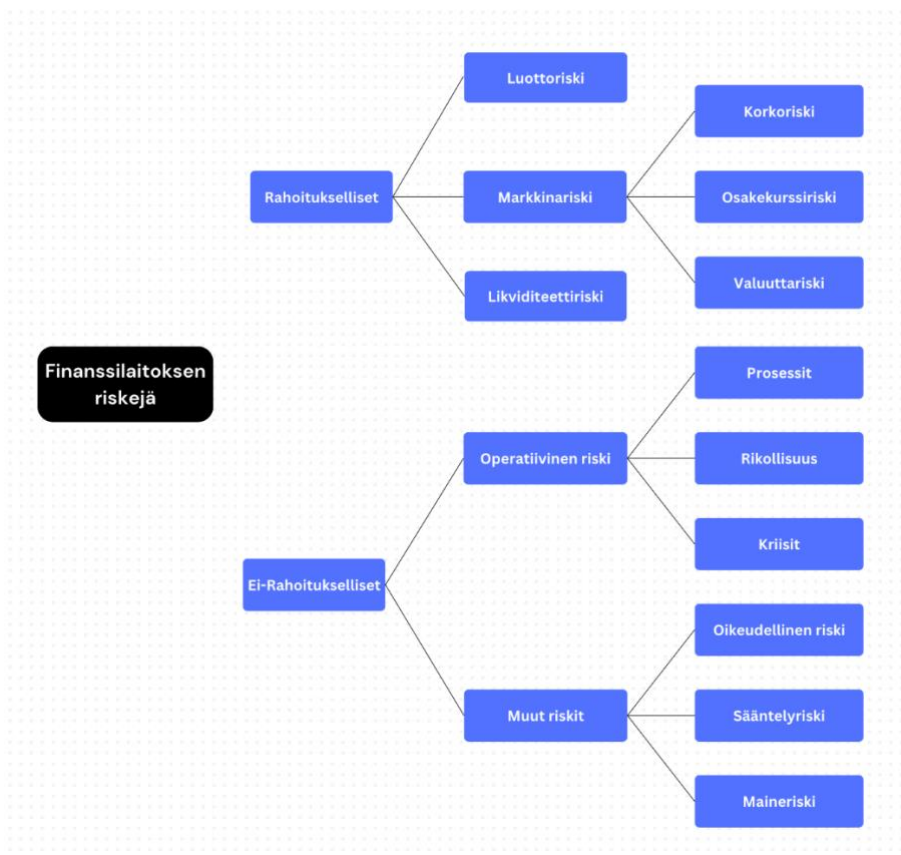


Kuva 2 Ennustavan mallin rakentamisen prosessi (mukaillen Nyce, 2007; Fawcett & Provost, 2013)

### 3 Ennakoivan analytiikan käyttö finanssilaitosten riskienhallinnassa

#### 3.1 Tyypillisiä finanssilaitosten riskejä

Finanssilaitokset joutuvat jatkuvasti arvioimaan ja hallitsemaan erilaisia riskejä. Ne voidaan jakaa rahoituksellisiin ja ei-rahoituksellisiin riskeihin (ks. Kuva 3). Ennen kuin perehdytään syvällisemmin erilaisiin riskeihin, on hyvä määrittää mitä riskillä tarkoitetaan. Holton (2004) määrittää riskin mitattavissa olevaksi epävarmuudeksi. Hansson (2013) taas puhuu tiedon puutteesta ja epäedullisesta lopputuloksesta (engl. unwanted event) määritellessään riskiä. Molempien määritelmistä löytyy yhtenäisyys tiedon puutteen ja epävarmuuden välillä. Etenkin Hansson (2013) toteaa ytimekkäästi, että jos tiedämme palavan talon räjähtävän varmasti, emme voi puhua sen räjähtämisen riskistä. Voidaan siis todeta, että riski syntyy tiedon puutteesta, joka johtaa epävarmuuteen tulevasta. Finanssisektorin kontekstissa tämän kyseisen epävarmuuden mittaaminen on riskienhallinnan keskiössä. Mitattavia asioita ovat epäsuotuisan tapahtuman vakavuus ja sen toteutumisen todennäköisyys (Kanchu & Kumar, 2013). Ennen kuin siirrytään erilaisten riskityyppien tarkempaan kuvaukseen, on hyvä todeta, että finanssilaitosten riskit ovat monitahoisia ja liittyvät toisiinsa.



Kuva 3 Finanssilaitoksen tyypillisiä riskejä (mukaiillen Leo ym., 2019)

Luottoriskillä tarkoitetaan riskiä, jonka finanssilaitos ottaa lainatessaan rahaa asiakkailleen. Riski realisoituu, kun lainanottaja ei täytä velvoitteitaan eli maksa lainaa ja sen kuluja takaisin. (Moges ym., 2013.) Finanssilaitokset, erityisesti pankit, altistuvat merkittävässä määrin luottoriskille, sillä luotonanto on niiden liiketoiminnan ydin. Luottoriski sisältää lainojen laiminlyöntien lisäksi myös lainanottajien luottokelpoisuuden heikkenemisen. Tällöin lainanantajan riski-tuottosuhte heikkenee. (Kanchu & Kumar, 2013.) Leo ym. (2019) lisäävät, että luottoriski on finanssilaitosten kokemista riskeistä merkittävin. Myös Van Greuning & Brajovic Bratanovic (2009, s. 161) toteavat luottoriskin olevan merkittävin riskityyppi. Heidän mukaansa jopa 70 % finanssilaitoksen taseesta liittyy tähän riskienhallinnan osa-alueeseen.

Markkinariskillä taas viitataan taloudellisten tappioiden riskiin, joka johtuu markkinoiden muutoksista (Wang, 2021). Se voidaan jakaa korkoriskiin, valuuttariskiin ja osakekurssiriskiin (Leo ym., 2019). Vashisht ym. (2022) kertovat, että markkinariski on erityisen olennainen kaupankäyntiä ja sijoittamista harjoittavissa instituutioissa, koska omistusten arvojen volatilitteetti vaikuttaa suoraan niiden taloudelliseen tulokseen.

Wang, (2021) väittää, että finanssilaitokset ovat kroonisesti herkkiä korkoriskille. Tämä johtuu siitä, että pankkien talletusten ja lainojen välinen suhde ei ole optimaalinen, koska talletussyklit ovat lyhyitä ja lainasyklit pitkiä. Jos korot laskevat niin pankit saavat pienempää katetta lainaamilleen varoilleen. (Wang, 2021.) Lisäksi finanssilaitosten liiketoiminnan laajentuminen perinteisestä luotonannosta kokonaisvaltaiseksi sijoitustuotteiden tarjoajaksi altistaa ne enenevässä määrin markkinariskille. Tämä johtuu siitä, että finanssilaitosten tulos on yhä riippuvaisempi markkinoiden kehityksestä, kuin perinteisestä luotonannosta tulevasta kassavirrasta. (Van Greuning & Brajovic Bratanovic, 2009, s. 228.) Markkinariskin realisoituessa sen vaikutus voi olla jopa 10 kertaa suurempi, kuin luottoriskin, erityisesti finanssikriisien aikana (Vashisht ym., 2022).

Likviditeettiriskin hallinta on keskeisimpiä tekijöitä maailmanlaajuisen rahoitusjärjestelmän luottamuksen rakentamisessa. Syynä on se, että finanssilaitokset ovat luontaisesti todella vivutettuja ja niiden omaisuuserien ja ydinpääoman suhde voi olla jopa 20/1. Suhde kuulostaa todella epälikvidiltä, mutta asianmukaisen riskienhallinnan avulla rahoitusjärjestelmän osapuolet voivat luottaa siihen, että omat varat ovat turvassa ja nostettavissa kun siihen on tarve. Likviditeetin asianmukaista hallintaa ei voi väheksyä, sillä yksittäisen finanssilaitoksen likviditeettiriskin realisoituminen saattaa aiheuttaa jälkiseurauksia koko järjestelmään. (Van Greuning & Brajovic Bratanovic, 2009, s. 193-194.)

Likviditeettiriski voidaan määritellä todennäköisyydeksi jolla finanssilaitos ei kykene täyttämään lyhyen aikavälin taloudellisia velvoitteitansa (Drehmann & Nikolaou, 2013). Tämä riski voi ilmetä, kun finanssilaitokseen kohdistuu äkillisiä nostoja, tai kun se ei kykene myymään omaisuuttaan nopeasti ilman merkittäviä tappioita. Likviditeettikriisit voivat syntyä sisäisesti, jolla viitataan asiakkaiden luottamuksen menettämiseen, tai ulkoisista rahoitussokeista. (Megeid, 2017.)

Vashishtin ym. (2022) mukaan vuoden 2008 finanssikriisi korosti likviditeetin hallinnan merkitystä, sillä monet finanssilaitokset ajautuivat maksukyvyttömyyteen riittämättömien likviditeettipuskureiden vuoksi. Asianmukaiset likviditeetin hallintastrategiat, kuten reservien ylläpitäminen ja rahoituslähteiden arviointi, ovat kriittisiä tämän riskin pienentämiseksi.

Rahoituksellisten riskien lisäksi finanssilaitokset joutuvat hallitsemaan ei-rahoituksellisia riskejä. Ei-rahoituksellisilla riskeillä tarkoitetaan riskejä, jotka realisoituessaan voivat vaikuttaa finanssilaitoksen liiketoiminnan kasvuun tai maineeseen. Merkittävin ei-rahoituksellinen riski on operatiivinen riski. (Kanchu & Kumar, 2013.)

Operatiivinen riski käsittää tappiot, jotka johtuvat puutteellisista tai epäonnistuneista sisäisistä prosesseista, järjestelmistä tai ulkoisista tapahtumista (Leo ym., 2019). Tällaisia voivat olla esimerkiksi inhimilliset virheet, tärkeän tietojärjestelmän rikkoutuminen, kyber- ja muut palvelunestohyökkäykset sekä luonnonkatastrofit ja kriisit (Kanchu & Kumar, 2013). Kun finanssilaitokset implementoivat yhä enemmän tietotekniikkaa osaksi prosessejaan, operatiivisen riskin todennäköisyys kasvaa. Lisäksi internetin välityksellä tapahtuva rikollisuus, kuten luottokorttivarkaudet, roskaposti (engl. spam) ja rahanpesu, on yleistynyt, mikä edelleen aiheuttaa haasteita finanssilaitoksille. (Leo ym., 2019; Van Greuning & Brajovic Bratanovic, 2009, s. 295.)

Uusien älykkäiden järjestelmien implementoinnin on huomattu pienentävän markkina- ja luottoriskiä, koska ne tarjoavat hyödyllistä informaatiota päätöksentekijälle (Van Greuning & Brajovic Bratanovic, 2009, s. 295). Finanssilaitosten riskienhallintapäälliköt joutuvatkin pohtimaan, että missä suhteessa uudet järjestelmät tarjoavat etua rahoituksellisten riskien optimoinnissa ja kuinka paljon ne altistavat operatiiviselle riskille. Van Greuning ja Brajovic Bratanovic (2009, s. 295) väittävät, että markkina- ja luottoriskiä voidaan vain optimoida, mutta operatiivista riskiä pyritään aina minimoimaan.

Operatiivinen riski ei ole ainoa ei-rahoituksellinen riski, jota finanssilaitokset joutuvat hallitsemaan. Muita ei-rahoituksellisia riskejä ovat muun muassa sääntelyriski, oikeudellinen riski ja maineriski (Leo ym., 2019). Reimersin ja Scheepersin (2016) kirjallisuuskatsauksesta selviää, että ei-rahoituksellisten riskityyppien luokittelu ei ole yksiselitteistä. Esimerkiksi Baselin

pankkivalvontakomitean (engl. Basel Committee on Banking Supervision, BCBS) määritelmässä oikeudellinen riski kuuluu operatiivisten riskien luokkaan. On hyvä ymmärtää, että riskityyppien luokittelu nähdään hanakalana myös finanssialan työntekijöiden keskuudessa, koska useat eri riskityypit ovat sidoksissa toisiinsa (Reimers & Scheepers 2016). Reimers ja Scheepers (2016) huomauttavat, että oleellista on hahmottaa riskit kokonaisuutena, eikä vain puretua yhteen luokkaan.

Sääntelyriskillä tarkoitetaan lakien ja määräysten noudattamatta jättämisestä mahdollisesti aiheutuvia oikeudellisia seuraamuksia, taloudellisia tappioita tai mainevahinkoja. Finanssilaitoksia säännellään voimakkaasti ja säännöksiä noudattamatta jättäminen voi johtaa sakkoihin, toimintarajoituksiin sekä mainevahinkoihin. (Losiewicz-Dniestrzanska, 2015.) Gündüz (2020) toteaa, että sääntelyriskin merkitys on kasvanut rahoitusmarkkinoiden monimutkaisuuden myötä.

Maineriski on riski, joka syntyy asiakkaiden tai muiden osapuolien kielteisistä käsityksistä finanssilaitosta kohtaan, mitkä voivat hankaloittaa sen liiketoimintaa, kuten uusien asiakkaiden hankkimista (Scandizzo, 2011). Maineriskin todennäköisyys on kasvanut digitaalisella aikakaudella, koska sosiaalisessa mediassa uutiset ja kuulopuheet voivat levitä yhdessä yössä miljoonille ihmisille. Maineriskin realisoituminen voi aiheuttaa asiakkaiden joukkopaon, jolloin useat asiakkaat nostavat varojaan pankista, mikä voi johtaa likviditeettiriskin realisoitumiseen ja edelleen jopa konkurssiin. (Hamidi ym., 2021.)

Yhteenvedona voidaan todeta, että finanssilaitosten kohtaamat riskit ovat moninaisia ja kytköksissä toisiinsa, mikä edellyttää ennakoivia hallintastrategioita. Luottoriskistä maineriskiin, jokainen luokka aiheuttaa ainutlaatuisia haasteita, jotka voivat vaikuttaa finanssilaitosten vakauteen ja kannattavuuteen. Kirjallisuudessa korostetaan, että näiden riskien tehokas käsittely edellyttää kehittyneen analytiikan, vankan sääntelyn noudattamisen ja jatkuvan valppauden yhdistelmää operatiivisissa prosesseissa (Javaid, 2024; Leo ym., 2019; Olaiya ym., 2024). Seuraavassa alaluvussa perehdytään tarkemmin ennakoivan analytiikan rooliin finanssilaitosten riskienhallinnassa.

### **3.2 Ennakoiva analytiikka finanssilaitosten riskienhallinnan apuvälineenä**

Finanssilaitokset kääntyvät enenevässä määrin ennakoivan analytiikan puoleen parantaakseen riskienhallintakykyään. Ennakoiva analytiikka hyödyntää valtavia määriä dataa ja kehittyneitä algoritmeja, mikä mahdollistaa ennakoivan lähestymistavan mahdollisten riskien hallintaan jatkuvasti kehittyvässä taloudellisessa maisemassa. (Javaid, 2024; Olaiya ym., 2024.) Tässä

alaluvussa käsitellään ennakoivan analytiikan hyödyntämistä finanssilaitosten riskienhallinnan eri osa-alueilla, ja verrataan näitä kehittyneitä tekniikoita tavanomaisempiin menetelmiin. Alaluvussa havainnollistetaan, mitä hyötyjä ja haittoja ennakoivien menetelmien hyödyntäminen tuo tullessaan.

Yksi ennakoivan analytiikan ensisijaisista sovelluksista finanssilaitoksissa on luottoriskin arviointi. Finanssilaitokset arvioivat lainanottajan maksukyvyttömyyden todennäköisyyttä (Javaid, 2024; Leo ym., 2019). Perinteisesti luottopisteitysmallit ovat perustuneet yksinkertaisempiin tekniikoihin, jotka kykenevät käsittelemään vain rajallisen määrän muuttujia lainanottajan luottokelpoisuuden arvioimiseksi. Ennakoiva analytiikka tarjoaa monimutkaisempia, esimerkiksi koneoppimis pohjaisia ennustavia malleja, jotka kykenevät analysoimaan useampia muuttujia tehokkaasti. (Javaid, 2024.) Näiden kehittyneiden mallien hyödyt eivät rajoitu vain niiden kykyyn hyödyntää suurempien tietomääriä. Kehittyneet ennustavat mallit kykenevät tunnistamaan heikompia suhteita ja trendejä selittävien muuttujien ja kohdemuuttujan välillä. (Olaiya ym., 2024.) Lisäksi koneoppimisalgoritmien avulla voidaan tunnistaa uusia selittäviä muuttujia, jotka vaikuttavat merkittävästi kohdemuuttujan arvoon (Fawcett & Provost, 2013, 48-50). Tämä auttaa finanssilaitoksia ymmärtämään luottoriskin juurisyitä. Tarkempi ennustava voima (engl. predictive power) antaa päätöksentekijälle mahdollisuuden tehdä parempia päätöksiä (Shmueli & Koppius, 2011).

Luokittelupuu- ja klusterointimalleilla lainanottajat voidaan jakaa ominaisuuksiensa mukaan eri kategorioihin (ks. Kuvio 1). Niiden avulla finanssilaitokset voivat hahmottaa asiakkaan riskiluokan ja hinnoitella tuotteensa tarkemmin. (Farayola ym., 2024.) Luokittelussa voidaan käyttää sääntöihin perustuvia tai itsestään oppivia malleja. Itsestään oppivat mallit ovat osoittautuneet tehokkaiksi, koska ne eivät pidä sisällään inhimillisiä vinoumia. Ne kategorisoivat muuttujia niiden ominaisuuksien perusteella tiettyihin kategorioihin ilman syvällisempiä perusteluja. (Broby, 2022.) Tämä voi toisinaan osoittautua myös ongelmaksi, jos tulosten perusteille olisi tarvetta. Tuntemalla asiakkaan riskiprofiilin ja hinnoitteleamalla tuotteet yksilöllisesti sen mukaan, finanssilaitokset kykenevät optimoimaan luottoriskiään ja pienentämään mahdollisia luottotappioita (Leo ym., 2019).

Ennakoiva analytiikka tarjoaa lukuisia etuja riskienhallintaan verrattuna perinteisiin menetelmiin, mutta mallien monimutkaisuus asettaa haasteita läpinäkyvyydelle ja tulkittavuudelle (Farayola ym., 2024). Perinteiset staattiset mallit ovat yleensä helpommin ymmärrettävissä ja perusteltavissa, mikä on olennaista tiukasti säännellyissä ympäristöissä. Vastaukseksi mallien monimutkaisuudelle on kehitetty selittävän tekoälyn (engl. explainable artificial intelligence, XAI) -tekniikoita, joiden avulla finanssilaitokset voivat tulkita ja perustella monimutkaisten ennustavien mallien tekemiä

päätöksiä (Bücker ym., 2022). Avoimuus on ratkaisevan tärkeää sidosryhmien luottamuksen rakentamisessa ja sääntelyvaatimusten noudattamisen varmistamisessa (Farayola ym., 2024). Esimerkiksi luottojen ja lainojen pisteytyksessä XAI voi selventää, miten tietyt tekijät vaikuttivat tulokseen. Tämä auttaa finanssilaitoksia pysymään laajan regulaation asettamien vaatimusten mukaisina.

Operatiivinen riski, joka sisältää järjestelmävirheistä, inhimillisistä virheistä tai ulkoisista tapahtumista johtuvat tappiot, on toinen riskienhallinnan osa-alue, jossa ennakoiva analytiikka on osoittautunut hyödylliseksi. (Leo ym., 2019.) Ennustavat mallit analysoivat menneitä tapahtumia tunnistaakseen trendejä, jotka voisivat viitata mahdollisiin toiminnallisiin epäonnistumisiin (Javaid, 2024). Esimerkiksi instituutiot voivat ennakoivan analytiikan avulla ennakoida järjestelmäkatkoksia seuraamalla järjestelmäinfrastruktuurin toimintatietoja ja puuttamalla ennakoivasti haavoittuvuuksiin (Farayola ym., 2024). Tämä ennakoiva lähestymistapa on vastakkainen perinteisille menetelmille, jotka ovat usein reaktiivisia ja käsittelevät asioita vasta niiden tapahtumisen jälkeen (Javaid, 2024; Olaiya ym., 2024). Ennakoivan analytiikan avulla finanssilaitokset voivat minimoida operatiivisia häiriöitä, mikä vähentää niihin liittyviä kustannuksia ja parantaa kokonaistehokkuutta (Farayola ym., 2024).

Kuitenkin merkittävin ennakoivan analytiikan sovellusalue operatiivisten riskien hallinnassa on petosten havaitseminen ja rikollisuuden torjunta (Leo ym., 2019). Perinteiset petosten havaitsemisjärjestelmät tukeutuvat staattisiin, sääntöihin perustuviin lähestymistapoihin. Tällaiset mallit sisältävät ennalta määritettyjä kynnyksarvoja ja järjestelmä huomauttaa, kun ne ylitetään. (Olaiya ym., 2024.) Hälytyksen aiheuttamia tapahtumia voivat olla esimerkiksi epätavallisen suuret siirrot tai tapahtumat odottamattomissa paikoissa. Vaikka nämä perinteiset menetelmät ovat jossain määrin tehokkaita, ongelmaksi muodostuu niille tyypillinen viive. Koneoppimis pohjaiset mallit hyödyntävät jatkuvasti uutta dataa ja reagoivat epäilyttäviin tapahtumiin reaaliajassa, mikä vaikeuttaa rikollisten toimintaa. (Javaid, 2024.) Samalla kun internetin välityksellä tapahtuva rikollisuus kasvaa, myös sen tekniikat kehittyvät. Tämä on kasvava haaste finanssilaitoksille. (Statista, 2024.)

Haasteeseen voidaan vastata ennakoivan analytiikan avulla. Yksi vastaus on reaaliaikaista dataa hyödyntävät luokittelualgoritmit, kuten klusterointi ja päätöspuut. Nämä mallit pystyvät havaitsemaan tapahtumia, jotka poikkeavat tyypillisestä transaktiokäyttäytymisestä. Reaaliaikainen analyysi auttaa havaitsemaan petoksia ennen merkittäviä vahinkoja. (Olaiya ym., 2024.)

Petosten, kuten luottokorttivarkauksien lisäksi rikollisuus, etenkin rahanpesu, on merkittävä riski finanssilaitoksille. Siinä missä petosten tekniikat ovat kehittyneet, myös rahanpesun havaitsemisesta on tullut haastavampaa. Rahanpesussa rikolliset pyrkivät peittelemään rahan alkuperää ketjuttamalla useita tilisiirtoja eri finanssilaitosten välillä. Transaktiot saattavat vaikuttaa täysin legitiimeiltä, jos ei ole ymmärrystä kokonaiskuvasta. Ennakoivan analytiikan avulla transaktioita kyetään tarkastelemaan laajemmassa kontekstissa. Esimerkiksi klusteroinnin avulla asiakkaat, joilla on samankaltaisuuksia transaktioissa, voidaan jakaa ryhmiin ja siten tunnistaa henkilöitä, jotka ovat mahdollisesti osana suurempaa rahanpesuvyyhtiä. (Leo ym., 2019.)

Operatiivisten riskien hallintaan liittyy myös suojautuminen kyberhyökkäyksiä ja tietomurtoja vastaan. Kyberhyökkäykset voivat kohdistua merkittäviin tietojärjestelmiin ja niiden tavoitteena on lamauttaa instituution toiminta. Tietomurroilla viitataan hyökkäyksiin, joiden tavoitteena on varastaa dataa, joka finanssilaitoksen kontekstissa voi tarkoittaa henkilökohtaisia tili- ja luottotietoja. (Darem ym., 2023.) Hyökkäykset saattavat esiintyä roskapostin muodossa. Se vaikuttaa negatiivisesti työntekijöiden tuottavuuteen. Tämän lisäksi roskapostia lähettävät tahot voivat olla kiinnostuneita yrityksen datasta, jota pyritään varastamaan haittaohjelmia (engl. malware) sisältävillä viesteillä. Koneoppimispohjaisilla ennustavilla malleilla voidaan suodattaa miljoonia viestejä reaaliajassa, joka auttaa yrityksiä suojautumaan haitalliselta roskapostilta. (Leo ym., 2019.)

Operatiivisten riskien hallinta on luottojen pisteytyksen lisäksi merkittävin ennakoivan analytiikan sovellusalue finanssilaitosten riskienhallinnassa. Petosten, rahanpesun, kyberhyökkäysten ja tietomurtojen havaitseminen on keskeinen haaste finanssilaitoksille, jotka operoivat jatkuvasti muuttuvassa ympäristössä. Näiden haasteiden torjunta edellyttää reaaliaikaisia valvontajärjestelmiä, jotka kykenevät käsittelemään valtavia määriä dataa. (Olaiya ym., 2024.)

Likviditeettiriski syntyy, kun finanssilaitos ei pysty täyttämään lyhyen aikavälin rahoitusvelvoitteitaan (Drehmann & Nikolaou, 2013). Ennakoiva analytiikka auttaa hallitsemaan tätä riskiä analysoimalla historiallisia kassavirtatietoja. Historiallisten kassavirtojen perusteella voidaan ennustaa tulevia tuloja, mikä auttaa finanssilaitosta sopeuttamaan toimintaansa ja varautumaan mahdollisiin taloudellisesti heikkoihin jaksoihin. (Broby, 2022.)

Myös markkinariskien hallinnassa voidaan hyödyntää ennakoivaa analytiikkaa. Esimerkiksi erilaisten kohde-etuksien volatiliteettia ja arvon muutoksia voidaan ennustaa ennustavilla malleilla. Ennustavat mallit tarjoavat päätöksen tekijälle informaatiota, jonka avulla hän voi muun muassa optimoida sijoitusportfolioita ja reagoida markkinoiden muutoksiin nopeasti. (Javaid, 2024.)

Ennakoiva analytiikka tarjoaa useita etuja perinteisiin riskienhallintamenetelmiin verrattuna. Seuraavaksi tutkielmassa vedetään yhteen keskeisimmät hyödyt ja haitat, jotka tulisi tunnistaa ennen ennakoivien riskienhallintatekniikoiden käyttöönottoa.

**Sopeutumiskyky ja oppiminen:** Ennakoivat analytiikkatekniikat, erityisesti koneoppiminen, ovat mukautuvia ja voivat parantua ajan myötä uutta dataa käsittelemällä (Javaid, 2024). Perinteiset menetelmät ovat staattisempia, ja ne nojaavat ennalta määriteltyihin sääntöihin. Tämä tekee niistä vähemmän responsiivisia, joka voi osoittautua ongelmaksi nopeasti muuttuvassa ympäristössä. (Olaiya ym., 2024.)

**Datan laajuus ja monipuolisuus:** Ennakoiva analytiikka hyödyntää sekä strukturoituja että strukturoimattomia tietolähteitä saadakseen kattavamman kuvan riskeistä (Olaiya ym., 2024). Sen sijaan tavanomaiset mallit hyödyntävät ensisijaisesti vain strukturoitua dataa, mikä rajoittaa niiden analyysin laajuutta. (Javaid, 2024.)

**Monimutkaisuus ja läpinäkyvyys:** Koneoppimispohjaiset ennustavat mallit ovat tehokkaampia, kuin perinteiset staattiset mallit, mutta ne vaativat kehittyneempää järjestelmäinfrastruktuuria ja ovat usein haastavia tulkita. Vastakkainasettelu staattisten ja itsestään oppivien mallien välillä kiteytyykin mallien ennustavan voiman ja tulkittavuuden ristiriitaan. Perinteiset menetelmät, vaikkakin yksinkertaisemmat, tarjoavat suuremman läpinäkyvyyden, mutta häviävät ennustusten tarkkuudessa. (Shmueli & Koppius, 2011.)

Yhteenvedona voidaan todeta, että ennakoiva analytiikka on muuttanut finanssilaitosten riskienhallintaa mahdollistamalla ennakoivan, datalähtöisen lähestymistavan (Javaid, 2024). Hyödyntämällä kehittyneitä tekniikoita luotto-, markkina-, likviditeetti- ja operatiivisten riskien hallinnassa, finanssilaitokset voivat ennakoida riskejä tehokkaammin ja optimoida päätöksentekoprosessejaan (Javaid, 2024; Leo ym., 2019; Olaiya ym., 2024). Ennustaviin malleihin liittyvät monimutkaisuus- ja sääntelyhaasteet edellyttävät huolellista toimintaa sekä läpinäkyvyyden, että vaatimustenmukaisuuden varmistamiseksi (Adebukola ym., 2022; Farayola ym., 2024). Selitettävän tekoälyn ja datarikkaiden mallien yhdistelmän avulla finanssilaitokset voivat hallita yhä monimutkaisempia riskejä ennakoivasti ja entistä tehokkaammin.

## 4 Yhteenveto ja johtopäätökset

Tutkielmassa selvitettiin miten finanssilaitokset hyödyntävät ennakoivaa analytiikkaa riskienhallinnassa. Tämän tavoitteen saavuttamiseksi esitettiin kaksi tutkimuskysymystä: Mitä on ennakoiva analytiikka ja miten sitä käytetään rahoituslaitosten riskienhallinnassa? Tässä luvussa vastataan näihin kysymyksiin ja esitetään tutkielman tieteellinen viitekehys ennakoivan analytiikan hyödyntämisestä finanssilaitosten riskienhallinnassa.

Taulukko 1 Tieteellinen viitekehys ennakoivan analytiikan käytöstä finanssilaitosten riskienhallinnassa

Käsite	Kuvaus	Esimerkkejä
Data	Laadukas, virheetön ja reaaliaikainen data on luotettavien ennustavien mallien perusta.	Asiakasdata, transaktiodata, markkinadata, strukturoimattomat datalähteet
Ennakoivan analytiikan tekniikat	Tekniikat datan analysointiin ja mallien rakentamiseen valitaan riskityypin perusteella.	Regressio (luottoriski), klusterointi ja luokittelupuut (petosten tunnistus), koneoppiminen
Tekniikoiden implementoiminen	Ennustavat mallit implementoidaan osaksi työprosesseja ja päätöksenteon tukijärjestelmiä	Luottojenpisteytys, petosten havaitseminen, rikollisuuden torjunta, portfolion optimointi
Eettiset näkökulmat	Läpinäkyvyyden ja oikeudenmukaisuuden varmistaminen	Vinoumien välttäminen luottojenpisteytyksessä, tietosuojavaatimusten noudattaminen datan hallinnassa

Ennakoiva analytiikka on ylätasoinen käsite kehittyneille tilastollisille tekniikoille, joiden avulla pyritään ennustamaan tulevia tapahtumia historiallisten tietojen perusteella. Sille on ominaista sen sopeutumiskyky, riippuvuus suurista tietoaaineistoista ja kyky paljastaa piileviä trendejä ja suhteita datassa. Se sisältää erilaisia tilastollisia- ja koneoppimis pohjaisia tekniikoita, kuten regressioanalyysi, klusterointi ja päätöspuut. Näiden menetelmien avulla finanssilaitokset voivat hyödyntää laajoja tietoaaineistoja ja muokata niitä hyödylliseksi informaatioksi. Tutkielmassa todettiin datan laadun ja sopivan menetelmän valinnan olevan oleellisia seikkoja tehokkaan ennustavan mallin rakentamisessa. Lisäksi mallin selitettävyyteen ja mahdollisiin vinoumiin tulee kiinnittää huomiota.

Menetelmän valinta riippuu käsillä olevan ongelman luonteesta. Tutkielmassa perehdyttiin finanssilaitosten tyypillisiin riskeihin ja todettiin, että erilaisten riskityyppien hallinta vaatii erilaisia lähestymistapoja. Tyypilliseksi riskeiksi luokiteltiin luotto-, markkina-, likviditeetti- ja operatiivinen riski. Jokainen näistä esittää ainutlaatuisia haasteita finanssilaitoksille.

Luottoriski on merkittävin finanssilaitosten kohtaama riski ja sen hallinta vaatii lainanottajan käyttäytymisen ja taloudellisten tekijöiden huolellista arviointia. Markkinarisktiin liittyy markkinoiden muutoksiin, jotka johtavat taloudellisiin tappioihin. Se voidaan edelleen jakaa korko-, valuutta- ja osakekurssiriskisiin. Likviditeettiriski syntyy, kun finanssilaitos ei pysty täyttämään lyhyen aikavälin rahoituksellisia velvoitteitaan. Operatiivinen riski syntyy prosessien epäonnistumisista, inhimillisistä virheistä tai ulkoisista tapahtumista, kuten kriiseistä tai kyberrikollisuudesta, joka on yhä yleisempää nykypäivän digitaalisessa ympäristössä. Näiden lisäksi maine-, sääntely- ja oikeudelliset riskit vaativat myös huomiota.

Tutkielman toinen tutkimuskysymys liittyy ennakoiva analytiikan rooliin edellä mainittujen riskien hallinnassa. Tutkielmassa todettiin, että ennakoiva analytiikka on osoittautunut päteväksi työkaluksi näiden riskien hallinnassa. Se mahdollistaa ennakoivan ja datalähtöisen päätöksenteon, jolloin riskienhallinnan luonne muuttuu reaktiivisesta proaktiiviseksi.

Luottoriskin hallinnassa koneoppimisalgoritmit parantavat lainanottajien arviointien tarkkuutta, kun taas operatiivisten riskien kohdalla ennakoivat mallit havaitsevat poikkeavuuksia, jotka voivat viestittää petoksista tai järjestelmäongelmista. Markkina- ja likviditeettiriskien hallintaan ennakoiva analytiikka tarjoaa tietoa markkinoiden muutoksista ja kassavirroista. Tämä auttaa finanssilaitoksia kohdentamaan resursseja ja ennakoimaan taloudellisesti heikkoja jaksoja proaktiivisesti.

Tämän tutkielman havainnot korostavat ennakoivan analytiikan potentiaalia finanssilaitosten riskienhallinnan mullistajana. Entistä proaktiivisempi ja tarkempi riskienhallinta parantaa finanssilaitosten resilienssiä ja lisää sidosryhmien luottamusta niitä kohtaan. Datan laatuun, mallien läpinäkyvyyteen ja säännösten noudattamiseen liittyviä haasteita ei kuitenkaan voi sivuuttaa. Tulevassa tutkimuksessa voitaisiin tutkia ennakoivan analytiikan integrointia kehittyviin teknologioihin, kuten lohkoketjuun, vastaamaan näihin haasteisiin.

## Lähteet

- Adebukola, A. A., Navya, A. N., Jordan, F. J., Jenifer, N. J., & Begley, R. D. (2022). Cyber Security as a Threat to Health Care. *Journal of Technology and Systems*, 4(1), 32–64. <https://doi.org/10.47941/jts.1149>
- Broby, D. (2022). The use of predictive analytics in finance. *The Journal of Finance and Data Science*, 8, 145–161. <https://doi.org/10.1016/j.jfds.2022.05.003>
- Bücker, M., Szepannek, G., Gosiewska, A., & Biecek, P. (2022). Transparency, auditability, and explainability of machine learning models in credit scoring. *Journal of the Operational Research Society*, 73(1), 70–90. <https://doi.org/10.1080/01605682.2021.1922098>
- Chopra, S. (2019). *Supply chain management: Strategy, planning and operation* (Seventh edition). Pearson Education.
- Darem, A. A., Alhashmi, A. A., Alkhaldi, T. M., Alashjaee, A. M., Alanazi, S. M., & Ebad, S. A. (2023). Cyber Threats Classifications and Countermeasures in Banking and Financial Sector. *IEEE Access*, 11, 125138–125158. <https://doi.org/10.1109/ACCESS.2023.3327016>
- Drehmann, M., & Nikolaou, K. (2013). Funding liquidity risk: Definition and measurement. *Journal of Banking & Finance*, 37(7), 2173–2182. <https://doi.org/10.1016/j.jbankfin.2012.01.002>
- Duggan, W. (2023, kesäkuuta 21). *A Short History Of The Great Recession*. Forbes Advisor. <https://www.forbes.com/advisor/investing/great-recession/>
- Farayola, O. A., Adaga, E. M., Egieya, Z. E., Ewuga, S. K., Abdul, A. A., & Abrahams, T. O. (2024). Advancements in predictive analytics: A philosophical and practical overview. *World Journal of Advanced Research and Reviews*, 21(3), 240–252. <https://doi.org/10.30574/wjarr.2024.21.3.2706>
- Fawcett, T., & Provost, F. (2013). *Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking*. O'Reilly Media, Inc.
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.
- Gündüz, V. (2020). Risk Management in Banking Sector. *Teoksessa Management & Strategy* (ss. 121–135). Artikel Akademi.
- Hamidi, S. R., Akmar Ismail, M., & Shuhidan, S. M. (2021). Corporate Reputation Risk in Social Media. *2021 International Conference on Computer Science and Engineering (IC2SE)*, 1, 1–8. <https://doi.org/10.1109/IC2SE52832.2021.9792037>
- Hansson, S. (2013). *The Ethics of Risk: Ethical Analysis in an Uncertain World*. Springer.

- Holton, G. A. (2004). Defining Risk. *Financial Analysts Journal*, 60(6), 19–25.  
<https://doi.org/10.2469/faj.v60.n6.2669>
- Infographic: *Cybercrime Expected To Skyrocket in Coming Years*. (2024, helmikuuta 22). Statista Daily Data. <https://www.statista.com/chart/28878/expected-cost-of-cybercrime-until-2027>
- Investointipankki *Lehman Brothers kaatui 10 vuotta sitten*. (2018, syyskuuta 15). Yle Uutiset.  
<https://yle.fi/a/3-10406455>
- Javaid, H. A. (2024). AI-Driven Predictive Analytics in Finance: Transforming Risk Assessment and Decision-Making. *Advances in Computer Sciences*, 7(1).  
<https://academicpinnacle.com/index.php/acs/article/view/204>
- Kalaian, S. A., & Kasim, R. M. (2017). Predictive Analytics. Teoksessa *Decision Management: Concepts, Methodologies, Tools, and Applications*. IGI Global. <https://doi.org/10.4018/978-1-5225-1837-2.ch004>
- Kanchu, T., & Kumar, M. M. (2013). Risk Management in Banking Sector—An Empirical Study. *International journal of Marketing, Financial Services & Management Research*, 2(2), 145–153.
- Leo, M., Sharma, S., & Maddulety, K. (2019). Machine Learning in Banking Risk Management: A Literature Review. *Risks*, 7(1), 29. <https://doi.org/10.3390/risks7010029>
- Losiewicz-Dniestrzanska, E. (2015). Monitoring of Compliance Risk in the Bank. *Procedia Economics and Finance*, 26, 800–805. [https://doi.org/10.1016/S2212-5671\(15\)00846-1](https://doi.org/10.1016/S2212-5671(15)00846-1)
- Megeid, N. S. A. (2017). Liquidity risk management: Conventional versus Islamic banking system in Egypt. *Journal of Islamic Accounting and Business Research*, 8(1), 100–128.  
<https://doi.org/10.1108/JIABR-05-2014-0018>
- Močarníková, K., & Greguš, M. (2020). Conceptualization of Predictive Analytics by Literature Review. Teoksessa N. Kryvinska & M. Greguš (Toim.), *Data-Centric Business and Applications: Lecture Notes on Data Engineering and Communications Technologies* (Vsk. 2, ss. 205–234). Springer, Cham. [https://doi.org/10.1007/978-3-030-19069-9\\_8](https://doi.org/10.1007/978-3-030-19069-9_8)
- Moges, H.-T., Dejaeger, K., Lemahieu, W., & Baesens, B. (2013). A multidimensional analysis of data quality for credit risk management: New insights and challenges. *Information & Management*, 50(1), 43–58. <https://doi.org/10.1016/j.im.2012.10.001>
- Nyce, C. (2007). *Predictive Analytics White Paper*. <https://www.the-digital-insurer.com/wp-content/uploads/2013/12/78-Predictive-Modeling-White-Paper.pdf>
- Olaiya, O., Agwubuo, C., Usoro, S., Obani, O., Nwafor, K., & Ajayi, O. (2024). The impact of big data analytics on financial risk management. *International Journal of Science and Research Archive*, 12, 821–827. <https://doi.org/10.30574/ijrsra.2024.12.2.1313>

- Ravi, K., Khandelwal, Y., Krishna, B. S., & Ravi, V. (2018). Analytics in/for cloud-an interdependence: A review. *Journal of Network and Computer Applications*, *102*, 17–37. <https://doi.org/10.1016/j.jnca.2017.11.006>
- Reimers, C., & Scheepers, C. B. (2016). Exploring the role of non-financial risk management in strategy processes of large retail banks. *South African Journal of Business Management*, *47*(3), Article 3.
- Salzberg, S. L. (1994). C4.5: Programs for Machine Learning by J. Ross Quinlan. Morgan Kaufmann Publishers, Inc., 1993. *Machine Learning*, *16*(3), 235–240. <https://doi.org/10.1007/BF00993309>
- Scandizzo, S. (2011). A framework for the analysis of reputational risk. *The Journal of Operational Risk*, *6*(3), 41–63.
- Shmueli, G., & Koppius, O. (2011). Predictive Analytics in Information Systems Research. *MIS Quarterly*, *35*, 553–572. <https://doi.org/10.2139/ssrn.1606674>
- Tummino, M. (2018). *Exploring the Use of Predictive Analytics in Banking and Finance Decision-Making*. [https://vc.bridgew.edu/honors\\_proj/262](https://vc.bridgew.edu/honors_proj/262)
- Van Greuning, H., & Brajovic Bratanovic, S. (2009). *Analyzing Banking Risk: A Framework for Assessing Corporate Governance and Risk Management, Third Edition* (3. p.). The World Bank. <https://doi.org/10.1596/978-0-8213-7728-4>
- Vashisht, S., Sarva, M., & Mundi, H. S. (2022). Risks measurement in banking: A bibliometric and content analysis. *International Social Science Journal*, *72*(246), 955–977. Scopus. <https://doi.org/10.1111/issj.12371>
- Wang, H. (2021). Credit Risk Management of Consumer Finance Based on Big Data. *Mobile Information Systems*, *2021*(1), 8189255. <https://doi.org/10.1155/2021/8189255>
- Zhang, D., Yin, C., Zeng, J., Yuan, X., & Zhang, P. (2020). Combining structured and unstructured data for predictive models: A deep learning approach. *BMC Medical Informatics and Decision Making*, *20*(1), 280. <https://doi.org/10.1186/s12911-020-01297-6>